Research White Paper

Observational Studies: Empirical Evidence of Their Contributions to Comparative Effectiveness Reviews

Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Research White Paper

Observational Studies: Empirical Evidence of Their Contributions to Comparative Effectiveness Reviews

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No.: 290-2013-00013-I

Prepared by:
Alberta Evidence-based Practice Center
Edmonton, Alberta, Canada

Investigators:
Jennifer C. Seida, M.P.H.
Donna M. Dryden, Ph.D.
Lisa Hartling, Ph.D.

AHRQ Publication No. 14-EHC002-EF
December 2013

This report is based on research conducted by the Alberta Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2013-00013-I). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

This report may periodically be assessed for the urgency to update. If an assessment is done, the resulting surveillance report describing the methodology and findings will be found on the Effective Health Care Program Web site at www.effectivehealthcare.ahrq.gov. Search on the title of the report.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact EffectiveHealthCare@ahrq.hhs.gov.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

Suggested citation: Seida JC, Dryden DM, Hartling L. Observational Studies: Empirical Evidence of Their Contributions to Comparative Effectiveness Reviews. Methods White Paper (Prepared by the University of Alberta Evidence-based Practice Center under Contract No. 290-2013-00013-I.) AHRQ Publication No. 14-EHC002-EF. Rockville, MD: Agency for Healthcare Research and Quality. December 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To improve the scientific rigor of these evidence reports, AHRQ supports empiric research by the EPCs to help understand or improve complex methodologic issues in systematic reviews. These methods research projects are intended to contribute to the research base in and be used to improve the science of systematic reviews. They are not intended to be guidance to the EPC program, although may be considered by EPCs along with other scientific research when determining EPC program methods guidance.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality. The reports undergo peer review prior to their release as a final report.

We welcome comments on this Methods Research Project. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to epc@ahrq.hhs.gov.

Richard G. Kronick, Ph.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang M.D., M.P.H.
Director EPC Program and Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Observational Studies: Empirical Evidence of Their Contributions to Comparative Effectiveness Reviews

Structured Abstract

Introduction. Although observational studies are increasingly being used to address gaps in the evidence from randomized controlled trials, the effect they have on the results and conclusions of systematic reviews is unclear. Our objectives were to evaluate: (1) how often observational studies are searched for and included in comparative effectiveness reviews (CERs); (2) the rationale for including or excluding observational studies; (3) how data from observational studies are appraised, analyzed, and graded; and (4) the impact of observational studies on the strength of evidence (SOE) and overall conclusions.

Methods. In June 2013 we searched the Effective Health Care Program Web site for final reports of CERs. One reviewer screened titles, abstracts, and Key Questions for CERs that examined a therapeutic or preventive intervention provided at an individual patient level. We selected a 25 percent sample of the most recent eligible CERs. Data were extracted by one reviewer and verified by a second reviewer. We extracted the number and type of study designs included and the approaches to quality assessment, presentation of results, and grading the SOE. We identified all comparisons for which both trials and observational studies provided data, and evaluated whether observational studies had an impact on the SOE and conclusions. We applied an RCT filter to the searches to determine the impact on search yield.

Results. From 129 records we identified 88 eligible CERs. Our final sample included 23 CERs published since November 2012. EPCs searched for observational studies in 20 CERs, of which 18 included a median of 11 (interquartile range: 2, 31) studies. Sixteen CERs incorporated the observational studies in their SOE assessments. We identified 78 comparisons from 12 CERs for which both trials and observational studies provided evidence; observational studies had an impact on SOE and conclusions for 19 (24 percent) of the comparisons. There was considerable diversity across the CERs regarding decisions to include or exclude observational studies, the study designs considered, and the approaches used to appraise, synthesize, and grade the SOE. Applying an RCT filter reduced the search yield by 65 percent (range 39 to 92 percent).

Discussion. Reporting guidelines and methods guidance relating to observational studies are needed in order to ensure clarity and consistency across Evidence-based Practice Centers. It was not always clear that the inclusion of observational studies added value in light of the additional resources needed to search for, select, appraise, and analyze such studies.

Contents

Introduction

Historically, systematic reviews of health care interventions have focused primarily on randomized controlled trials (RCTs) because they are more likely to provide unbiased data about the differential effects of treatment options than other study designs.[1] Because randomization controls for both known and unknown confounders, differences between the study groups of RCTs can be attributed to the intervention of interest. However, for some interventions, randomization may not be considered ethical or feasible. In addition, trials may provide inadequate data on harms due to short followup periods or have limited applicability to real-world settings. For some comparisons of interest, evidence from RCTs may be sparse or unavailable.

Systematic reviewers are increasingly considering evidence from observational studies to address gaps in the RCT evidence.[2,3] Observational studies may provide longer followup data, be conducted in settings that more closely resemble clinical practice, report on rare outcomes, and provide evidence for interventions that cannot be randomized.[1] Although there is general consensus that observational studies, especially those based on large clinical or administrative databases, provide a valuable contribution for assessing potential harms, the role of observational studies in addressing questions of clinical benefit is much disputed.[4] Concerns about the risks of bias inherent to nonexperimental studies, particularly due to baseline differences between the intervention groups, are often cited as reasons to exclude observational studies from systematic reviews.[1,5] Furthermore, the inclusion of observational studies may also add significant time and resource requirements to the review process.

Several organizations have provided guidance on when to consider including observational studies in systematic reviews. In their Standards for Systematic Reviews, the Institute of Medicine recommends that observational studies be used in addition to RCTs to evaluate harms of interventions and be considered to address gaps in the evidence from RCTs on the benefits of interventions.[6] The Cochrane Handbook stipulates that the inclusion of nonrandomized studies may be justified to provide: (a) an evaluation of their methodological weaknesses in the context of building a case for the need for RCTs; (b) evidence of the effects of interventions that cannot be randomized; and c) evidence of effects that cannot be adequately examined in RCTs.[1] Similarly, guidance from the Grading of Recommendations Assessment, Development and Evaluation Working Group suggests that observational studies may be used to complement data from trials or when no trials are available.[7]

The Agency for Healthcare Research and Quality (AHRQ) recently developed a conceptual framework for considering the inclusion of observational studies in comparative effectiveness reviews (CERs) of therapeutic interventions.[8] For CERs, which compare the relative benefits and harms among a range of available interventions for a given condition, reviewers are cautioned against developing protocols that rule out the inclusion of observational studies a priori. RCTs are thought to often be insufficient to address all aspects of CER research questions.[8] Guidelines for deciding whether or not to use evidence from observational studies include whether gaps exist in the evidence from RCTs and whether observational studies provide useful and valid information.[8]

Despite guidance that supports considering observational studies for inclusion in systematic reviews, there is a lack of empirical evidence to demonstrate the impact observational studies have on their findings, strength of evidence (SOE) assessments, and conclusions. Because the inclusion of observational studies involves additional workload and may increase resource requirements and timelines, it is important to understand when and in what context (e.g.,

different topics, clinical areas) observational studies are likely to provide additional evidence and strengthen the conclusions of a CER. An Evidence-based Practice Center (EPC) Working Group is concurrently preparing a white paper that presents case examples of when observational studies were used in evidence reviews of therapeutic interventions and how they contributed to the overall evidence synthesis. Our methods report expands on this work by generating empirical evidence of the impact observational studies have on the results and conclusions of CERs from the EPC program.

Our objectives were to systematically review CERs of therapeutic and preventive intervention to identify: (1) how often observational studies are searched for and included; (2) the rationale for including or excluding observational studies; (3) how data from observational studies are assessed for validity, presented, analyzed, and incorporated in the SOE assessments and conclusions; and (4) the impact or contribution of observational studies on the results, SOE, and overall conclusions.

Methods

We conducted a study of a sample of recently published CERs that have been conducted within the EPC program. To identify CERs, we searched the Effective Health Care Program Web site for all final reports of research reviews published in English. The search was conducted in June 2013. A single reviewer screened the titles, abstracts, and Key Questions of these reviews by applying predefined inclusion criteria. In order to be eligible for our study, the review was required to be the final report of a CER and have at least one Key Question that: (a) examined a therapeutic or preventive intervention (i.e., screening and diagnostic reviews were excluded); and (b) provided an intervention at an individual-patient level (i.e., program or health systems interventions were excluded). We selected a 25 percent sample of the eligible CERs, based on most recent date of publication, due to time constraints. The most recent CERs were selected because they are more representative of current EPC practices, particularly with respect to SOE assessments.

The data extraction form is available in Appendix A. Data were extracted by one reviewer using a standard form and verified by a second reviewer. Discrepancies were resolved by discussion or third-party arbitration.

First, we extracted descriptive data from each CER, including the clinical area, type of intervention (e.g., drug, surgery, nonpharmacological), year of publication, total number of studies included, and the number and types of Key Questions (e.g., clinical benefit, adverse events). We extracted information on whether or not the reviewers intended to include observational studies; the rationale provided for the decision to include or exclude observational studies (e.g., long-term followup, safety outcomes, a priori impression that there are few RCTs on a topic); and the types of observational studies eligible for inclusion in the CER.

We defined observational studies as any nonexperimental design in which the investigators did not assign group exposure. Observational studies could be prospective or retrospective in their design or analysis, and could have a concurrent, historical, or within-group comparison or no comparison group. Common observational designs include cohort, case-control, before-after, cross-sectional, case series, and case report studies. As there is variability in the terminology used to describe study designs, we used the classification system developed by the Alberta EPC.[9]

For Key Questions that included observational studies, we extracted the number and type of study designs included, the tools used to assess study quality, and the approaches used to present study findings. Specifically, we extracted information on whether the observational data were analyzed quantitatively in a meta-analysis (either separately or together with trials), qualitatively (e.g., narrative description of findings), or displayed graphically in a visual analysis. We looked for examples of CERs that presented data in particularly effective ways in order to generate ideas on novel approaches to summarizing findings from observational studies. We recorded whether observational studies were included in the SOE assessments and any specific methods described for grading observational studies (e.g., additional domains assessed, how assessments for observational studies were combined with those for trials).

For outcomes that were graded, we identified all comparisons for which both trials and observational studies provided data. We focused on these comparisons due to resource limitations. Using the SOE tables, we assessed whether observational studies appeared to have an impact on the individual domains (risk of bias, consistency, directness, and precision) and on the overall SOE assessments. We rated whether observational studies had an impact on SOE as "yes," "no," or "unclear" for each comparison and outcome. When SOE tables were not available in the report or appendixes, or the number of studies and the designs that contributed to

each comparison and outcome were not reported, we rated the impact of observational studies as "unclear." The assessments were made by one reviewer and verified by a second reviewer. For each comparison and outcome, we recorded comments justifying our assessments.

We developed principles to guide our assessments of the impact of observational studies through an iterative process and discussions about the various scenarios encountered. For each outcome, we considered the proportion of studies and patients contributed by observational studies versus trials; the quality of the observational studies compared with the trials; whether the findings from observational studies were consistent with those of the trials; whether there were differences in the directness of the comparisons or outcomes between observational studies and trials; and whether precision appeared to be affected by the inclusion of observational studies. When more than one trial reported on a given outcome, we assessed that the impact of observational studies was minimal unless the findings were discordant with those observed in the trials. When the overall SOE was rated insufficient, we generally assessed that the observational studies had made no contribution to the SOE unless it was apparent that the addition of observational studies lowered the SOE in one or more domains.

We engaged in a similar iterative process to evaluate the impact of observational studies on CER conclusions. When observational studies were evaluated to have no impact on SOE, we rated that they also did not affect the conclusions unless the conclusions incorporated the balance between benefits and harms and observational studies contributed to the harms evidence. When the conclusions reiterated the SOE rating, then we provided the same rating across both categories. We considered conclusions reported in the abstract, executive summary, and the discussion section of the CER because we found that sometimes different outcomes were included or different wording was used within individual reports. Further, we conducted subgroup analyses to examine whether any characteristics of the CERs (e.g., health condition, intervention type, types of observational studies included) were associated with a more likely impact of observational studies on SOE assessments.

Finally, we re-ran the Medline or PubMed search strategies with and without a Cochrane RCT filter for a sample of five CERs conducted by different EPCs across a variety of topics. We compared the yield of records identified through the two searches to estimate the added time and human resources involved in including observational studies alongside trials in a CER.

Results

Our search for final reports of CERs on the Effective Health Care Program Web site yielded 129 records. Of the 129 reports, we identified 88 eligible CERs (76 of a therapeutic intervention, 12 of a preventive intervention). From the eligible CERs, we initially selected a sample of 25 percent of the most recently published CERs to review (n=20 therapeutic, n=3 preventive). A list of the CERs included in our study is available in Appendix B. Table 1 provides an overview of the characteristics of the CERs.

The 23 CERs were conducted across nine EPCs and published between November 2012 and June 2013. They focused on a range of health condition categories: brain or nerve (n=2), breathing (n=2), cancer (n=4), eye, ear or throat (n=1), physical disabilities (n=1), genitourinary (n=1), gynecology (n=1), heart and blood vessel (n=4), infectious disease (n=2), mental health (n=3), and obesity (n=2). The majority of the CERs examined pharmaceutical interventions (n=18), either alongside other types of nonpharmaceutical interventions (n=15) or as the only intervention category (n=3). Surgical (n=7), procedural (n=7), and behavioral or educational (n=9) interventions were also commonly examined.

The reviewers from the EPCs searched for observational studies in 20 CERs (87 percent), including 2 of the 3 preventive CERs. Only seven CERs (30 percent) provided a rationale for their decision to include or exclude observational studies. Common reasons cited to justify including observational studies were to provide additional data for adverse events (n=4), a limited number of trials (n=1), and to target special populations that are typically excluded from trials (n=1). Reasons given to justify excluding observational studies were the methodological limitations of observational studies (n=2) and a sufficient number of trials available (n=1).

The observational study designs that were most often considered eligible were prospective or retrospective cohort studies (n=20) and case control studies (n=9). Case series and case reports were eligible in four reviews. Several CERs considered all observational studies to be eligible (n=4) or excluded only case reports (n=4). Three CERs used a hierarchy of evidence approach, such as including noncomparative observational studies only if comparative observational studies were not identified.

Half of the CERs (n=11) used different study design criteria for different Key Questions. Some reviewers decided a priori that it would be inappropriate to base decisions about effectiveness on observational studies, and therefore excluded these studies from Key Questions that focused on clinical benefit.[10] Overall, observational studies were eligible for inclusion in 68 percent of the Key Questions examining only clinical benefits, 73 percent of Key Questions examining both clinical benefits and harms, and 85 percent of Key Questions examining only harms.

A median of 11 observational studies (interquartile range: 2, 31) were included in the 20 CERs that searched for observational studies. Two CERs intended to include observational studies, but found no studies to meet their eligibility criteria.[10,11] The total number of included observational studies was unclear in three CERs. Table 1 provides the number of observational studies included in each CER by study design. Overall, the CERs predominantly included cohort studies and case series. In some reviews, study designs were difficult to categorize due to ambiguity of terms, such as “retrospective review,” “nonrandomized comparative study,” and “uncontrolled cohort.” Occasionally, the distribution of study designs could not be determined because all studies were described simply as “trials” or “observational.” Some CERs used study design algorithms, such as the taxonomy developed by the Alberta EPC, to categorize the included studies.[12,13]

A wide range of approaches were used to assess the quality of observational studies. The quality assessment tools included the Newcastle-Ottawa Scale (n=3), Cochrane Risk of Bias tool (n=3), Carey and Boden (n=2), Downs and Black scale (n=1), and the United States Preventive Services Task Force (USPSTF) tool (n=1). Some CERs reported rating quality items identified by the EPC (n=4) or stated that the approach to quality assessment was based on the EPC methods guide (n=6). Four CERs reported using different quality assessment tools for different observational study designs. In six CERs, a quality threshold was used to determine which observational studies would contribute data to the review.[10,12-16] For example, some reviews excluded studies rated to be at high risk of bias from the results or SOE assessment.[13,14]

The majority of the CERs provided a narrative description of the results of observational studies (n=16). Observational studies were included in a meta-analysis in three CERs;[17-19] in two, the data from observational studies were pooled with trial data.[17,19] One CER plotted the results of observational studies on a graph without conducting a meta-analysis.[19]

Of the 18 CERs that included observational studies, 16 incorporated observational studies in grading the SOE. Only seven CERs provided some description of the methods used to grade the observational studies.[18-24] Evidence from observational studies was graded together with trials in seven CERs and separately from trials in six; in three CERs it was unclear whether SOE was graded separately or together with trials. Observational studies were not included in the SOE assessments in two CERs.

We identified 78 comparisons from 12 CERs for which both trials and observational studies provided evidence and whose outcomes were graded. Observational studies had an impact on the SOE assessments for 19 comparisons (24 percent), had no impact for 37 comparisons (47 percent), and had an unclear impact for 22 comparisons (28 percent). Similarly, observational studies were assessed to have had an impact on the conclusions of the CER for 18 comparisons (23 percent), no impact for 39 comparisons (50 percent), and an unclear impact for 21 comparisons (27 percent).

We conducted subgroup analyses to determine whether the impact of observational studies on SOE assessments and conclusions was related to various characteristics of the CERs. Our subgroup analyses were limited by the small number of CERs (n=12) that provided data. Four CERs examined heart and blood vessel conditions, two examined brain and nerve conditions, and the remaining six CERs each examined a different health condition. Observational studies had an effect on the SOE in 35 percent of the comparisons on heart and blood vessel CER topics compared with 15 percent of all other CER health topics combined. Similarly, observational studies appeared to have a greater impact in comparisons that examined a surgical intervention (31 percent; n=26 comparisons) and devices (33 percent; n=9 comparisons) than for pharmaceuticals (11 percent; n=18 comparisons). For CERs that only included comparative observational designs (i.e., cohort or case control studies), observational studies were more likely to have an impact on SOE assessments (46 percent; n=26 comparisons) compared with CERs that included comparative and uncontrolled designs (17 percent; n=42 comparisons).

We re-ran the Medline or PubMed searches for a sample of five CERs both without a study design filter and with the Cochrane RCT filter applied. Appendix C provides a summary table of the findings. On average, applying an RCT filter reduced the search yield by 65 percent or 2,031 records. Using an estimated average screening rate of 100 records per hour, an RCT filter may reduce screening time by approximately 20 hours. The yield reduction varied considerably across CERs, ranging from a 39 percent reduction to a 92 percent reduction, depending on the initial search strategy developed by the EPC.

Table 1. Characteristics of included comparative effectiveness reviews

CER Number, Author, Date, EPC	Condition and Intervention	Observational Studies	Included Studies	Observational Study QA, Analysis, and SOE	Eligible SOE Comparisons*
76: Chou R et al., Nov 2012, Oregon	Infectious Disease: Hepatitis C Therapeutic Pharmaceutical	Not eligible	—	—	No
82: Maglione MA et al., Jun 2013, Southern California	Obesity Therapeutic Pharmaceutical, Surgical, Behavioral/educational	All observational study designs with >3 patients eligible	4 trials 20 observational (10 cohort, 1 CC, 9 CS)	NR Narrative description Graded separately, then combined with trials	Yes
88: Penson DF et al., Dec 2012, Vanderbilt	Genitourinary: Cryptorchidism Therapeutic Pharmaceutical, Surgical	All observational studies excluding CRs	16 trials 24 observational (3 PCS, 21 RCS)	Newcastle-Ottawa Scale MA, Narrative description Graded separately, then combined with trials	Yes
89: Goldman FJ et al., Apr 2013, North Carolina	Mental Health: Trauma Therapeutic Pharmaceutical, Behavioral/educational	All observational study designs with >10 patients eligible	24 trials 1 cohort	EPC generated tool with items specific to study design Narrative description, studies with high ROB excluded from results Graded separately, then combined with trials	No
90: Saha S et al., May 2013, Oregon	Functional Limitations: Pressure Ulcers Therapeutic Pharmaceutical, Surgical, Procedural, Device/equipment	Different observational designs eligible based on available level of evidence per KQ	143 trials 31 observational (1 PCS, 2 RCS, 2 cohort NOS, 1 before-after, 7 CS, 18 NR)	EPC generated tool based on Downs and Black Narrative description Graded together with trials	Yes
91: Sun X et al., Dec 2012, Oregon	Infectious Disease: Hepatitis C Therapeutic Pharmaceutical, Tx adherence, Management of AEs	All study designs eligible	6 trials 6 observational (3 PCS, 3 RCS)	Newcastle-Ottawa Scale Narrative description Graded together with trials	Yes
92: Jonas DE et al., Apr 2013, RTI–UNC	Mental Health: Posttraumatic stress disorder Therapeutic Pharmaceutical, Behavioral/educational	Observational studies excluded for KQ 1–5, included for KQ6 with sample size >500	92 trials 0 observational	–	No

Table 1. Characteristics of included comparative effectiveness reviews (continued)

CER Number, Author, Date, EPC	Condition and Intervention	Observational Studies	Included Studies	Observational Study QA, Analysis, and SOE	Eligible SOE Comparisons*
93: Belinson S et al., Dec 2012, Blue Cross	Cancer: Colorectal Cancer Therapeutic Pharmaceutical, Surgical, Procedural	All study designs eligible	0 trials (1 RCT included, but treated as a CS) 30 CS (13 prospective, 17 retrospective)	Carey and Boden criteria Narrative description Graded separately, then combined with trials	No
94: Ferluga ED et al., Mar 2013, Vanderbilt	Brain or Nerve Condition: Cerebral Palsy Therapeutic Surgical, Behavioral/educational, Nutritional interventions	All study designs eligible except CRs as long as before-after data provided	1 trial 11 observational (1 PCS, 10 CS)	Newcastle-Ottawa Scale, AHRQ methods guide approach Narrative description Graded together with trials	Yes
96: Hartmann KE et al., Mar 2013, Vanderbilt	Gynecology: Abnormal Uterine Bleeding Therapeutic Pharmaceutical, Behavioral/educational, Device, CAM	All study designs eligible	0 trials 25 observational (cohort, CC, post-marketing surveillance studies, registry data (N not specified))	QA method not specified Narrative description –	No
100: McCrory DC et al., Jan 2013, Duke	Airway Condition: Chronic Cough Therapeutic Pharmaceutical, Behavioral/educational, CAM	RCTs and cohort studies eligible	45 trials 3 cohorts	AHRQ methods guide approach – –	No
101: Berkman ND et al., May 2013, RTI–UNC	ENT: Otitis Media with Effusion Therapeutic Pharmaceutical, Surgical, Procedural, CAM	Trials and observational studies eligible (PCS, RCS, CC only)	55 trials 4 observational (3 RCS, 1 not specified)	Modified Cochrane ROB tool Narrative description, studies with high ROB included only for harms Graded together with trials	Yes
107: Forman-Hoffman V et al., Feb 2013, RTI–UNC	Mental Health: Trauma Therapeutic Pharmaceutical, Behavioral/educational, CAM	All study designs eligible, studies with high ROB excluded	21 trials 1 PCS	Cochrane ROB tool, AHRQ criteria, RTI item bank Narrative description Graded separately, then combined with trials	No
108: Shamliyan TA et al., Jun 2013, Minnesota	Brain or Nerve Condition: Migraine Prevention Pharmaceutical, Non-pharmaceutical	Eligible study designs not specified	26 trials 16 other (4 retrospective reviews, 2 prospective reviews,1 chart review, 1 CS, 8 unspecified)	RTI item bank Narrative description SOE methods unclear, study designs maybe combined	Yes

Table 1. Characteristics of included comparative effectiveness reviews (continued)

CER Number, Author, Date, EPC	Condition and Intervention	Observational Studies	Included Studies	Observational Study QA, Analysis, and SOE	Eligible SOE Comparisons*
111: Lin SY et al., Mar 2013, Johns Hopkins	Airway Condition: Allergic Rhinoconjunctivits, Asthma Therapeutic Allergen-specific immunotherapy	Not eligible	–	–	No
112: Ratko TA et al., Jun 2013, Blue Cross	Cancer: Lung Cancer Therapeutic Procedure	All study designs eligible	5 trials 50 observational (single arm studies [15 prospective, 21 retrospective, 13 not specified], 1 retrospective nonrandomized comparative study)	Carey and Boden criteria Narrative description Graded separately, then combined with trials	No
113: Grant MD et al., Apr 2013, Blue Cross	Cancer: Anemia Therapeutic Pharmaceutical	Comparative observational studies >250 patients for KQ1	54 trials 0 observational	–	
114: Belinson S et al., May 2013, Blue Cross	Cancer: Hepatocellular Carcinoma Therapeutic Procedure	All study designs eligible, CRs included only for KQ on harms	6 trials 42 observational (1 PCS, 4 RCS, 2 CC, 32 CS, 3 CRs)	USPSTF for RCTs and nonrandomized comparative studies, Carey and Boden criteria for single-arm studies Narrative description Graded together with trials	
115: Wang Y et al., Jun 2013, Johns Hopkins	Obesity: Childhood Obesity Prevention Behavioral/educational, Community-based intervention	Not eligible	–	–	

Table 1. Characteristics of included comparative effectiveness reviews (continued)

CER Number, Author, Date, EPC	Condition and Intervention	Observational Studies	Included Studies	Observational Study QA, Analysis, and SOE	Eligible SOE Comparisons*
116: Singh S et al., May 2013, Johns Hopkins	Heart and Blood Vessel Condition: Venous Thromboembolism Prevention Pharmaceutical, Mechanical	All study designs eligible	6 trials 95 observational (16 PCS, 24 RCS, 6 CS, 8 CRs, 37 uncontrolled cohort, 4 not specified)	Downs and Black MA included both trials and OS, Narrative description Graded together with trials	
117: McCrory DC et al., Apr 2013, Duke	Heart and Blood Vessel Condition: Pulmonary Arterial Hypertension Therapeutic Pharmaceutical	Prospective, retrospective studies, registry data eligible for KQ3 only	28 trials 9 observational (4 retrospective CS, 5 prospective CS)	Cochrane ROB tool – SOE method unclear	
118: Jones WS et al., May 2013, Duke	Heart and Blood Vessel Condition: Peripheral Artery Disease Therapeutic Pharmaceutical, Surgical, Procedural, Behavioral/education	PCS, RCS eligible	40 trials 43 observational (3 PCS, 1 PCS, 39 unclear)	EPC methods guide MA included both trials and OS, Narrative description Graded together with trials	Yes
119: Al-Khatib SM et al., Jun 2013, Duke	Heart and Blood Vessel Condition: Atrial Fibrillation Therapeutic Pharmaceutical, Procedural	Prospective, retrospective studies, registries >100 patients	146 trials 2 observational (not specified)	Modified AHRQ methods guide Narrative description SOE method unclear	Yes

Notes: AEs = adverse events; CAM = complementary and alternative medicine; CC = case control; CER = comparative effectiveness review; CS = case series; CR = case report; KQ = Key Question; MA = meta-analysis; N = number; NOS = not otherwise specified; NR = not reported; PCS = prospective cohort study; QA = quality assessment; RCS = retrospective cohort study; RCT = randomized controlled trial; ROB = risk of bias; SOE = strength of evidence; Tx = treatment

*Comparison must be graded and have data from at least 1 trial and 1 observational study

Discussion

In our sample of 23 CERs—all published since November 2012—EPCs searched for observational studies in 20, of which 18 included observational studies. A median of 11 (interquartile range: 2, 31) observational studies were included. Of these 18 CERs, 16 incorporated the observational studies in their SOE assessments. We identified 78 comparisons from 12 CERs for which both trials and observational studies provided evidence. We determined that observational studies had an impact on SOE and conclusions for 19 (24 percent) of the comparisons.

We found considerable diversity across the CERs regarding: decisions to include or exclude observational studies; the observational study designs that were considered; and, whether to specify study designs a priori or use a step-wise approach. The rationale for decisions was rarely provided. Further, there was diversity across the CERs in the approaches used to appraise, synthesize, and grade the SOE of observational studies.

The terminology used for observational study designs was inconsistent not only across CERs, but also within individual reports. It was sometimes difficult to determine what features were meant by certain labels. For example, a "controlled comparative study" could refer to a controlled clinical trial (nonrandomized) or a prospective cohort study. In some CERs specific study designs were not reported and studies were simply described as "observational" or "nonrandomized." The terms "observational" and "nonrandomized" were sometimes used interchangeably although it was unclear whether controlled clinical trials or quasi-experimental study designs were considered for inclusion. By virtue of the study design, we may have more confidence in the results from a nonrandomized controlled trial than from a cohort study.

The quality of observational studies was assessed using a range of approaches and tools. Some EPCs used one tool across all study designs (e.g., USPSTF tool), while others used design-specific tools such as the Newcastle-Ottawa Scale for cohort studies. When observational studies were included in SOE assessments, it was not always clear how EPCs made decisions regarding the risk of bias domain when different quality assessment or risk of bias tools were used. It is challenging to derive an overall risk of bias assessment when different study designs are included and when different tools are used. For example, when there are two RCTs at high risk of bias and two cohort studies that are low risk of bias, what is the risk of bias for this cluster of studies that provide evidence for a specific outcome?

In assessing the SOE, some SOE tables did not indicate how many studies and which study designs contributed to the assessment. In some instances the numbers reported in the SOE table and the text of the CER did not match. This made it difficult to determine the impact of observational studies on the SOE and conclusions presented in the CERs.

For some CERs the SOE was graded separately for trials and observational studies, while others combined study designs. Where study designs were graded separately it was difficult to determine an overall SOE for an outcome. However, when the evidence for a comparison consists of several RCTs at low risk of bias and observational studies at high risk of bias, it may be appropriate to focus on the SOE assessment of trials alone and not combine study designs.

Next Steps

In light of the observed diversity in the inclusion and reporting of observational studies, we feel that reporting guidelines and methods guidance are warranted in order to better ensure clarity and consistency across CERs.

Reporting Guidelines

1. Clear reporting of the rationale for including or excluding observational studies for each Key Question;
2. Clear reporting of the type of observational studies that are considered for inclusion;
3. Consistent use of study design nomenclature (e.g., the Alberta EPC system) or study design features (e.g., the Cochrane Collaboration);
4. Clear reporting regarding how nonrandomized trials and observational studies are assessed for quality, analyzed, and graded;
5. Clear reporting of the number of studies of which designs are included in the CER;
6. Clear reporting of the number and designs of studies contributing evidence for comparisons on SOE tables.

Methods Guidance

The recent AHRQ report on grading the SOE[25] will clarify some issues regarding how to incorporate observational studies. However, there is still little guidance on how to combine the assessment of experimental and observational studies, and its absence will likely result in inconsistency across EPCs. The following are important questions that we believe need to be addressed through methods guidance.

1. When should observational studies be included in CERs and in SOE assessments?
2. Is there a hierarchy of study designs to consider when moving beyond RCTs?
3. Is it appropriate to set a sample size limit for observational studies (i.e., >1,000 patients, >10 patients)?
4. How do we assess risk of bias domain for SOE when different study designs use different tools?
5. What proportion of studies must be low risk of bias to give an overall score of "low"?
6. In what situations can observational studies affect SOE (either to raise or lower an assessment)?
7. How do we assess precision when observational studies are not combined in a meta-analysis?
8. How do we assess publication bias for observational studies?

Limitations

It was challenging to determine what the impact of observational studies was on the SOE or conclusions. It was difficult to decide which conclusions to evaluate since the conclusions often reiterated the SOE assessments. Rarely did CERs present a conclusion that combined the balance of benefits and harms (e.g., benefit is good but the risk of harms from observational studies is sufficient to temper conclusions). Finally the subgroup analyses are based on a small sample size. While the subgroups were specified a priori, the results should still be considered as hypothesis-generating.

Summary and Conclusion

In this paper, we examined a sample of 23 recent CERs to explore how observational studies contribute to the results and conclusions of CERs. We found considerable diversity across the CERs regarding decisions to include observational studies and the specific study designs that were considered. The rationale for decisions was rarely provided. We also observed variability in the approaches used to appraise, synthesize, and grade the SOE of observational studies.

It was not always clear that the inclusion of observational studies added value to the additional resources needed to search for, select, appraise, and analyze such studies. Reporting guidelines and methods guidance relating to observational studies is needed in order to ensure clarity and consistency in how observational studies are handled within the EPC Program.

References

1. Higgins J, Green S, eds. Cochrane handbook for systematic reviews of interventions Version 5.1.0 [updated March 2011]. 2011.
2. Moja L, Telaro E, D'Amico R, et al. Assessment of methodological quality of primary studies by systematic reviews: results of the metaquality cross sectional study. BMJ. 2005;330:1053-7.
3. Norris S, Atkins D. Challenges in using nonrandomized studies in systematic reviews of treatment interventions. Ann Intern Med. 2005;142:1112-9.
4. Chou R, Aronson N, Atkins D, et al. AHRQ series paper 4: Assessing harms when comparing medical interventions: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2010;63(5):502-12.
5. Vandenbroucke J. When are observational studies as credible as randomised trials? Lancet. 2004;363:1728-31.
6. Institute of Medicine. Finding what works in health care: Standards for systematic reviews. Washington, D.C.: Institute of Medicine; 2011.
7. The GRADE Working Group. GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol. 2011;64:1311-6.
8. Norris S, Atkins D, Bruening W, et al. Observational studies in systemic reviews of comparative effectiveness: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2011;64(11):1178-86.
9. Hartling L, Bond K, Harvey K, et al. Developing and testing a tool for the classification of study designs in systematic reviews of interventions and exposures: methods research report. AHRQ Publication No. 11-EHC-007.Rockville, MD: Agency for Healthcare Research and Quality; 2010.
10. Jonas DE, Cusack K, Forneris CA,et al. Psychological and Pharmacological Treatments for Adults With Posttraumatic Stress Disorder (PTSD). Comparative Effectiveness Review No. 92. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13-EHC011-EF. Rockville, MD: Agency for Healthcare Research and Quality. April 2013.
11. Grant MD, Piper M, Bohlius J, et al. Epoetin and Darbepoetin for Managing Anemia in Patients Undergoing Cancer Treatment: Comparative Effectiveness Update. Comparative Effectiveness Review No. 113. (Prepared by the Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center under Contract No. 290-2007-10058-I.) AHRQ Publication No. 13-EHC077-EF. Rockville, MD: Agency for Healthcare Research and Quality. April 2013.
12. Goldman Fraser J, Lloyd SW, Murphy RA, et al. Child Exposure to Trauma: Comparative Effectiveness of Interventions Addressing Maltreatment. Comparative Effectiveness Review No. 89. (Prepared by the RTIUNC Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13-EHC002-EF. Rockville, MD: Agency for Healthcare Research and Quality. April 2013.
13. Forman-Hoffman V, Knauer S, McKeeman J, et al. Child and Adolescent Exposure to Trauma: Comparative Effectiveness of Interventions Addressing Trauma Other Than Maltreatment or Family Violence. Comparative Effectiveness Review No. F. (Prepared by the RTI International-University of North Carolina at Chapel Hill Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13-EHC054-EF. Rockville, MD: Agency for Healthcare Research and Quality. February 2013.

14. Berkman ND, Wallace IF, Steiner MJ, et al. Otitis Media With Effusion: Comparative Effectiveness of Treatments. Comparative Effectiveness Review No. 101. (Prepared by the RTI-UNC Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13-EHC091-EF. Rockville, MD: Agency for Healthcare Research and Quality. May 2013.

15. Hartmann KE, Jerome RN, Lindegren ML, et al. Primary Care Management of Abnormal Uterine Bleeding. Comparative Effectiveness Review No. 96. (Prepared by the Vanderbilt Evidence-based Practice Center under Contract No. 290-2007-10065 I.) AHRQ Publication No. 13-EHC025-EF. Rockville, MD: Agency for Healthcare Research and Quality. March 2013.

16. Saha S, Smith MEB, Totten A, et al. Pressure Ulcer Treatment Strategies: Comparative Effectiveness. Comparative Effectiveness Review No. 90. (Prepared by the Oregon Evidence-based Practice Center under Contract No. 290-2007-10057-I.) AHRQ Publication No. 13-EHC003-EF. Rockville, MD: Agency for Healthcare Research and Quality. May 2013.

17. Jones WS, Schmit KM, Vemulapalli S, et al. Treatment Strategies for Patients With Peripheral Artery Disease. Comparative Effectiveness Review No. 118. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2007-10066-I.) AHRQ Publication No. 13-EHC090-EF. Rockville, MD: Agency for Healthcare Research and Quality. May 2013.

18. Penson DF, Krishnaswami S, Jules A, et al. Evaluation and Treatment of Cryptorchidism. Comparative Effectiveness Review No. 88. (Prepared by the Vanderbilt Evidence-based Practice Center under Contract No. 290-2007-10065-I.) AHRQ Publication No. 13-EHC001-EF. Rockville, MD: Agency for Healthcare Research and Quality. December 2012.

19. Singh S, Haut ER, Brotman DJ, et al. Pharmacologic and Mechanical Prophylaxis of Venous Thromboembolism Among Special Populations. Comparative Effectiveness Review No. 116. (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. 290-2007-10061-I.) AHRQ Publication No. 13-EHC082-1. Rockville, MD: Agency for Healthcare Research and Quality. May 2013.

20. Al-Khatib SM, Allen Lapointe N, Chatterjee R, et al. Treatment of Atrial Fibrillation. Comparative Effectiveness Review 119. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2007-10066-I.) AHRQ Publication No.13-EHC095-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2013.

21. Belinson S, Chopra R, Yang Y, et al. Local Hepatic Therapies for Metastases to the Liver From Unresectable Colorectal Cancer. Comparative Effectiveness Review No. 93. (Prepared by Blue Cross and Blue Cross Blue Shield Association Technology Evaluation Center under Contract No. 290-2007-10058-I.) AHRQ Publication No. 13-EHC014-EF. Rockville, MD: Agency for Healthcare Research and Quality. December 2012.

22. Belinson S, Yang Y, Chopra R, et al. Local Therapies for Unresectable Primary Hepatocellular Carcinoma. Comparative Effectiveness Review No. 114. (Prepared by the Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center under Contract No. 290-2007-10058-I.) AHRQ Publication No. 13-EHC069-EF. Rockville, MD: Agency for Healthcare Research and Quality. May 2013.

23. Ratko TA, Vats V, Brock J, et al. Local Nonsurgical Therapies for Stage I and Symptomatic Obstructive Non–Small-Cell Lung Cancer. Comparative Effectiveness Review No. 112. (Prepared by Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center under Contract No. 290-2007-10058-I.) AHRQ Publication No. 13-EHC071-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2013.

24. Shamliyan TA, Kane RL, Ramakrishnan R, et al. Migraine in Children: Preventive Pharmacologic Treatments. Comparative Effectiveness Review No. 108. (Prepared by the University of Minnesota Evidence-based Practice Center under Contract No. 290-2007-10064-I.) AHRQ Publication No. 13-EHC065-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2013.

25. Berkman ND, Wallace IF, Steiner MJ, et al. Otitis Media With Effusion: Comparative Effectiveness of Treatments. Comparative Effectiveness Review No. 101. (Prepared by the RTI-UNC Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13-EHC091-EF. Rockville, MD: Agency for Healthcare Research and Quality. May 2013

Appendix A. Data Extraction Form

I. Cover Page

CER no.:	First author (e.g., Lin MA):	Date (e.g., 2013-06):
EPC (e.g., Alberta):	AHRQ publication no.:	
CER Title:		
DE initials:	DV initials:	

II. Description of CER

Health condition category (see EHCP Web site):	Specific condition (e.g., asthma):
Intervention (check all that apply): ☐ Preventive ☐ Therapeutic	Intervention type (check all that apply): ☐ Pharmacological ☐ Surgical ☐ Procedural ☐ Behavioral/Educational/Psychological ☐ Other, specify:

III. Observational Studies: Eligibility Criteria

For any of the KQs which meet our eligibility criteria (see Table below):

Did the reviewers intend to include observational studies (OS)?	
Did the reviewers provide a rationale for including or excluding OS?	
What rationale was provided? ☐ NA *Include:* ☐ Long-term follow-up ☐ Safety / adverse event outcomes ☐ Limited trials available ☐ Comprehensiveness *Exclude:* ☐ Sufficient evidence from trials ☐ Methodological weaknesses of OS ☐ Other*, specify:*	Which OS designs were eligible for inclusion? ☐ None ☐ Prospective cohort ☐ Retrospective cohort ☐ Cohort, not otherwise defined ☐ Case control ☐ Interrupted time series ☐ Before-after / pre-post ☐ Cross-sectional ☐ Case series ☐ Case report ☐ Not reported / unclear ☐ Other, specify:
Was a hierarchy approach to inclusion used? If so, explain in Notes	Did inclusion of OS vary by Key Question?
Notes:	

KQ	Eligible?*	KQ category	Brief description of KQ	OS considered?	Notes: OS designs considered
1					
2					
3					
4					
5					
6					
Notes:					

KQ = Key Question; OS = observational study

*Key Questions that examine a therapeutic or preventive intervention given at an individual patient level are eligible; for Key Questions that are not eligible, do not complete remaining columns.

For CERs that did NOT include observational studies, DO NOT complete the remainder of this form.

IV. Observational Studies: Assessment and Analysis

No. included studies in eligible KQ *only*:

trials (RCTs or NRCTs) + OS = total included studies

How many of the following OS designs were included in the *eligible KQs*?

- Prospective cohort
- Retrospective cohort
- Cohort, not otherwise defined
- Case control
- Interrupted time series
- Before-after / pre-post
- Cross-sectional
- Case series
- Case report
- Not reported / unclear
- Other, specify:

Which tool(s) were used to assess the quality of OS?

- ☐ None; quality was not assessed.
- ☐ Newcastle-Ottawa
- ☐ Jadad
- ☐ Cochrane Risk of Bias
- ☐ Downs and Black
- ☐ Other, specify:

Were different quality assessment tools used for different OS designs?

If so, explain:

Was a quality threshold used as an inclusion criterion?

Did the reviewers conduct meta-analysis with data from OS?

If so, were data from OS pooled together with data from trials?

Is there a narrative description of the findings of the OS?

Is there a visual presentation (not meta-graph) of the results of OS?

If so, describe the type of graph or visual display of the results (include page no.):

Notes:

V. Strength of Evidence

Was the strength of the evidence (SOE) (i.e., GRADE) assessed in the CER?
Were OS included in the SOE assessment(s)?
Did the reviewers describe any methods specific to grading SOE for OS? If so, briefly summarize these methods:
Notes:

VI. Results and Conclusions

Complete the table for data when <u>all</u> the following criteria are met: a) eligible KQ; b) both trial and OS provide data for a comparison; and c) outcomes are graded. If no SOE table is available, select "unclear."

Eligible KQ	Comparison	Outcome	OS impact SOE/results?*	OS impact conclusions?*	Comments
Notes:					

*would results and/or conclusions have been any different if OS had not been included?

Appendix B. Included Comparative Effectiveness Reviews

Below is an alphabetical listing of the 23 comparative effectiveness reviews (CERs) (20 therapeutic, 3 preventive) examined in our empirical study. All of the reviews are available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Al-Khatib SM, Allen Lapointe N, Chatterjee R, Crowley MJ, Dupre ME, Kong DF, Lopes RD, Povsic TJ, Raju SS, Shah BR, Kosinski A, McBroom AJ, Chobot MM, Gray R, Sanders GD. Treatment of Atrial Fibrillation. Comparative Effectiveness Review 119. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2007-10066-I.) AHRQ Publication No.13-EHC095-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2013.

Belinson S, Chopra R, Yang Y, Shankaran V, Aronson N. Local Hepatic Therapies for Metastases to the Liver From Unresectable Colorectal Cancer. Comparative Effectiveness Review No. 93. (Prepared by Blue Cross and Blue Cross Blue Shield Association Technology Evaluation Center under Contract No. 290-2007-10058-I.) AHRQ Publication No. 13-EHC014-EF. Rockville, MD: Agency for Healthcare Research and Quality. December 2012.

Belinson S, Yang Y, Chopra R, Shankaran V, Samson D, Aronson N. Local Therapies for Unresectable Primary Hepatocellular Carcinoma. Comparative Effectiveness Review No. 114. (Prepared by the Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center under Contract No. 290-2007-10058-I.) AHRQ Publication No. 13-EHC069-EF. Rockville, MD: Agency for Healthcare Research and Quality. May 2013.

Berkman ND, Wallace IF, Steiner MJ, Harrison M, Greenblatt AM, Lohr KN, Kimple A, Yuen A. Otitis Media With Effusion: Comparative Effectiveness of Treatments. Comparative Effectiveness Review No. 101. (Prepared by the RTI-UNC Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13-EHC091-EF. Rockville, MD: Agency for Healthcare Research and Quality. May 2013.

Chou R, Hartung D, Rahman B, Wasson N, Cottrell E, Fu R. Treatment for Hepatitis C Virus Infection in Adults. Comparative Effectiveness Review No. 76. (Prepared by Oregon Evidence-based Practice Center under Contract No. 290-2007-10057-I.) AHRQ Publication No. 12(13)-EHC113-EF. Rockville, MD: Agency for Healthcare Research and Quality. November 2012.

Ferluga ED, Archer KR, Sathe NA, Krishnaswami S, Klint A, Lindegren ML, McPheeters ML. Interventions for Feeding and Nutrition in Cerebral Palsy. Comparative Effectiveness Review No. 94. (Prepared by the Vanderbilt Evidence-based Practice Center under Contract No. 290-2007-10065-I) AHRQ Publication No. 13-EHC015-EF. Rockville, MD: Agency for Healthcare Research and Quality. March 2013.

Forman-Hoffman V, Knauer S, McKeeman J, Zolotor A, Blanco R, Lloyd S, Tant E, Viswanathan M. Child and Adolescent Exposure to Trauma: Comparative Effectiveness of Interventions Addressing Trauma Other Than Maltreatment or Family Violence. Comparative Effectiveness Review No. F. (Prepared by the RTI International-University of North Carolina at Chapel Hill Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13-EHC054-EF. Rockville, MD: Agency for Healthcare Research and Quality. February 2013.

Goldman Fraser J, Lloyd SW, Murphy RA, Crowson MM, Casanueva C, Zolotor A, Coker-Schwimmer M, Letourneau K, Gilbert A, Swinson Evans T, Crotty K, Viswanathan M. Child Exposure to Trauma: Comparative Effectiveness of Interventions Addressing Maltreatment. Comparative Effectiveness Review No. 89. (Prepared by the RTIUNC Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13-EHC002-EF. Rockville, MD: Agency for Healthcare Research and Quality. April 2013.

Grant MD, Piper M, Bohlius J, Tonia T, Robert N, Vats V, Bonnell C, Ziegler KM, Aronson N. Epoetin and Darbepoetin for Managing Anemia in Patients Undergoing Cancer Treatment: Comparative Effectiveness Update. Comparative Effectiveness Review No. 113. (Prepared by the Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center under Contract No. 290-2007-10058-I.) AHRQ Publication No. 13-EHC077-EF. Rockville, MD: Agency for Healthcare Research and Quality. April 2013.

Hartmann KE, Jerome RN, Lindegren ML, Potter SA, Shields TC, Surawicz TS, Andrews JC. Primary Care Management of Abnormal Uterine Bleeding. Comparative Effectiveness Review No. 96. (Prepared by the Vanderbilt Evidence-based Practice Center under Contract No. 290-2007-10065 I.) AHRQ Publication No. 13-EHC025-EF. Rockville, MD: Agency for Healthcare Research and Quality. March 2013.

Jonas DE, Cusack K, Forneris CA, Wilkins TM, Sonis J, Middleton JC, Feltner C, Meredith D, Cavanaugh J, Brownley KA, Olmsted KR, Greenblatt A, Weil A, Gaynes BN. Psychological and Pharmacological Treatments for Adults With Posttraumatic Stress Disorder (PTSD). Comparative Effectiveness Review No. 92. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13-EHC011-EF. Rockville, MD: Agency for Healthcare Research and Quality. April 2013.

Jones WS, Schmit KM, Vemulapalli S, Subherwal S, Patel MR, Hasselblad V, Heidenfelder BL, Chobot MM, Posey R, Wing L, Sanders GD, Dolor RJ. Treatment Strategies for Patients With Peripheral Artery Disease. Comparative Effectiveness Review No. 118. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2007-10066-I.) AHRQ Publication No. 13-EHC090-EF. Rockville, MD: Agency for Healthcare Research and Quality. May 2013.

Lin SY, Erekosima N, Suarez-Cuervo C, Ramanathan M, Kim JM, Ward D, Chelladurai Y, Segal JB. Allergen-Specific Immunotherapy for the Treatment of Allergic Rhinoconjunctivitis and/or Asthma: Comparative Effectiveness Review. Comparative Effectiveness Review No. 111.(Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. 290-2007-10061-I.) AHRQ Publication No. 13-EHC061-EF. Rockville, MD: Agency for Healthcare Research and Quality. March 2013. Errata added May 2013.

Maglione MA, Maggard Gibbons M, Livhits M, Ewing B, Hu J, Ruelaz Maher A, Li Z, Perry T, Shekelle PG. Bariatric Surgery and Nonsurgical Therapy in Adults With Metabolic Conditions and a Body Mass Index of 30.0 to 34.9 kg/m². Comparative Effectiveness Review No. 82. (Prepared by the Southern California Evidence-based Practice Center under Contract No. 290-2007-10062-I.) AHRQ Publication No. 12(13)-EHC139-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2013.

McCrory DC, Coeytaux RR, Schmit KM, Kraft B, Kosinski AS, Mingo AM, Vann LM, Gilstrap DL, Hargett CW, Lugogo NL, Heidenfelder BL, Posey R, Irvine RJ, Wing L, Pendergast K, Dolor RJ. Pulmonary Arterial Hypertension: Screening, Management, and Treatment. Comparative Effectiveness Review No. 117. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2007-10066-I.) AHRQ Publication No. 13-EHC087-EF. Rockville, MD: Agency for Healthcare Research and Quality. April 2013.

McCrory DC, Coeytaux RR, Yancy WS Jr., Schmit KM, Kemper AR, Goode A, Hasselblad V, Heidenfelder BL, Irvine RJ, Musty MD, Gray R, Sanders GD. Assessment and Management of Chronic Cough. Comparative Effectiveness Review No. 100. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2007-10066-I.) AHRQ Publication No. 13-EHC032-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2013.

Penson DF, Krishnaswami S, Jules A, Seroogy JC, McPheeters ML. Evaluation and Treatment of Cryptorchidism. Comparative Effectiveness Review No. 88. (Prepared by the Vanderbilt Evidence-based Practice Center under Contract No. 290-2007-10065-I.) AHRQ Publication No. 13-EHC001-EF. Rockville, MD: Agency for Healthcare Research and Quality. December 2012.

Ratko TA, Vats V, Brock J, Ruffner BW Jr, Aronson N. Local Nonsurgical Therapies for Stage I and Symptomatic Obstructive Non–Small-Cell Lung Cancer. Comparative Effectiveness Review No. 112. (Prepared by Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center under Contract No. 290-2007-10058-I.) AHRQ Publication No. 13-EHC071-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2013.

Saha S, Smith MEB, Totten A, Fu R, Wasson N, Rahman B, Motu'apuaka M, Hickam DH. Pressure Ulcer Treatment Strategies: Comparative Effectiveness. Comparative Effectiveness Review No. 90. (Prepared by the Oregon Evidence-based Practice Center under Contract No. 290-2007-10057-I.) AHRQ Publication No. 13-EHC003-EF. Rockville, MD: Agency for Healthcare Research and Quality. May 2013.

Shamliyan TA, Kane RL, Ramakrishnan R, Taylor FR. Migraine in Children: Preventive Pharmacologic Treatments. Comparative Effectiveness Review No. 108. (Prepared by the University of Minnesota Evidence-based Practice Center under Contract No. 290-2007-10064-I.) AHRQ Publication No. 13-EHC065-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2013.

Singh S, Haut ER, Brotman DJ, Sharma R, Chelladurai Y, Shermock KM, Kebede S, Stevens KA, Prakasa KR, Shihab HM, Akande TO, Zeidan AM, Garcia LJ, Segal JB. Pharmacologic and Mechanical Prophylaxis of Venous Thromboembolism Among Special Populations. Comparative Effectiveness Review No. 116. (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. 290-2007-10061-I.) AHRQ Publication No. 13-EHC082-1. Rockville, MD: Agency for Healthcare Research and Quality. May 2013.

Sun X, Patnode CD, Williams C, Senger CA, Kapka TJ, Whitlock EP. Interventions To Improve Patient Adherence to Hepatitis C Treatment: Comparative Effectiveness. Comparative Effectiveness Review No. 91. (Prepared by the Oregon Evidence based Practice Center under Contract No. 290-2007-10057-I.) AHRQ Publication No. 13-EHC009-EF. Rockville, MD: Agency for Healthcare Research and Quality. December 2012.

Wang Y, Wu Y, Wilson RF, Bleich S, Cheskin L, Weston C, Showell N, Fawole O, Lau B, Segal J. Childhood Obesity Prevention Programs: Comparative Effectiveness Review and Meta-Analysis. Comparative Effectiveness Review No. 115. (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. 290-2007-10061-I.) AHRQ Publication No. 13-EHC081-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2013.

Appendix C. Search Yields With and Without an RCT Filter

Below is a summary of the search yields both with and without applying an RCT filter for a sample of five CERs.

Summary of Results:

Search	Description	Results
(1) Non–Small-Cell Lung Cancer	Original search (without RCT filter) Filtered search (Cochrane RCT filter applied)	3,386 1,121
(2) Migraine in Children	Original search (without RCT filter) Filtered search (Cochrane RCT filter applied)	680 319
(3) Pressure Ulcers	Original search (without RCT filter) Filtered search (Cochrane RCT filter applied)	5,038 424
(4) Otitis Media With Effusion	Original search (without SD filters) Search results (with original SD filters) Filtered search (Cochrane RCT filter applied)	7,252 3,170 882
(5) Body Mass Index of 30.0 to 34.9 kg/m²	Original search (without SD filters) Search results (with original SD filters) Filtered search (Cochrane RCT filter applied)	5,047 1,628 999